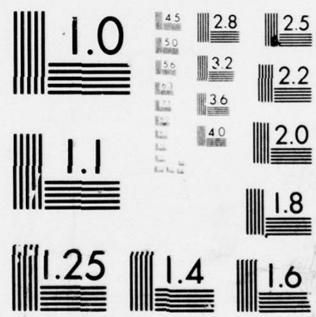
AD-A059 867 CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER F/G 6/16
ANALYSIS OF LIFE TABLES WITH GROUPING AND WITHDRAWALS.(U)
AUG 78 D V LINDLEY AFOSR-77-3179
UNCLASSIFIED ORC-78-13 NL

1 OF 1
AD
A059867

END
DATE
FILMED
12-78
DDC




MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12 LEVEL II

ORC 78-13
AUGUST 1978

# ANALYSIS OF LIFE TABLES WITH GROUPING AND WITHDRAWALS

by
DENNIS V. LINDLEY



OPERATIONS
RESEARCH
CENTER

UNIVERSITY OF CALIFORNIA · BERKELEY

ANALYSIS OF LIFE TABLES WITH GROUPING AND WITHDRAWALS[†]

by

Dennis V. Lindley[††]



AUGUST 1978



ORC 78-13

---

[†]This research was supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-77-3179 while the author was a Visiting Professor at the Operations Research Center, University of California, Berkeley. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[††]Address for correspondence: University College London, Gower Street, WC1E 6BT, London, England.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>ORC-78-13 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br>ANALYSIS OF LIFE TABLES WITH GROUPING AND WITHDRAWALS | | 5. TYPE OF REPORT & PERIOD COVERED<br>Research Report |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>Dennis V. Lindley | | 8. CONTRACT OR GRANT NUMBER(s)<br>AFOSR-77-3179 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Operations Research Center<br>University of California<br>Berkeley, California 94720 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS<br>2304/A5 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>United States Air Force<br>Air Force Office of Scientific Research<br>Bolling AFB, D.C. 20332 | | 12. REPORT DATE<br>August 1978 |
| | | 13. NUMBER OF PAGES<br>19 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)<br>12 20p. | | 15. SECURITY CLASS. (of this report)<br>Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Life Tables
Identifiability
Consistency
Bayesian Statistics

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(SEE ABSTRACT)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

270 750

ABSTRACT

Individuals are observed until either they withdraw or die. The data are then grouped into intervals and from the grouped data it is desired to estimate the death-rate for the intervals. The present paper considers only a single interval. A method is suggested which gives interesting estimates with standard errors that do not tend to zero as the sample size tends to infinity.

-A-

# ANALYSIS OF LIFE TABLES WITH GROUPING AND WITHDRAWALS

by

Dennis V. Lindley

We use the model of Breslow and Crowley (1974). Associated with each individual is a pair of independent random quantities, $X$, the lifetime and $Y$, the withdrawal time. The raw observations for an individual are $Z = \min(X,Y)$, the time at which he leaves either through death or withdrawal, and an indicator which says whether the departure was caused by death or withdrawal. The time scale is then divided into nonoverlapping intervals and the quantity $Z$ grouped so that observation is only made on the interval within which he left the system. The quantities for different individuals are independent and identically distributed. Consequently if $N$ individuals are present at the beginning of an interval, the data consists of $D$, the number who were observed to die in the interval; $W$ the number observed to withdraw alive during the interval; and $S$ the number who survived to enter the next interval. $N = D + W + S$. In this paper we shall only consider a single interval, except for a remark in the discussion at the end of the paper. Our task is to estimate the chance of death occurring in the interval. If $X$ and $Y$ have distribution functions $F$ and $H$ respectively and the interval is $(0,1]$ conventionally, then this chance is $F(1)$. The novelty lies in the fact that the probability of being observed to die is necessarily less than $F(1)$, since some deaths during the interval will take place after withdrawal and therefore be unobserved. This happens whenever $Y < X \leq 1$. The unknown quantity of interest is $F(1)$: the data are $D$, $W$ and $S$, so that we require to calculate $p(F(1) \mid D,W,S)$.

Another possible interpretation of the model is to replace "death" by "death from a specific cause," say cancer; and "withdrawal" by "death from other causes." In that situation we have a competing risk problem and both death-rates, $F(1)$ and $H(1)$, will be of interest. We shall therefore concentrate on the calculation of $p(F(1),H(1) \mid D,W,S)$ from which the earlier quantity can be obtained by taking the marginal distrubution. Notice that this competing-risk model assumes the two risks, expressed through $X$ and $Y$, are independent. The model has applications to life-tables in biology and in reliability engineering.

The model is specified in terms of $F(x)$ and $H(x)$ for $0 \leq x \leq 1$, with $F(0) = H(0) = 0$, so that it is natural to extend the conversation from the quantities of interest, $F(1)$ and $H(1)$, to include these functions. We therefore begin by calculating the likelihood of $F(x)$, $H(x)$, $0 \leq x \leq 1$ for the data $(D,W,S)$. An individual will be observed to die in the interval if he dies at any time $x$, $0 < x \leq 1$, and his withdrawal quantity $Y$ exceeds $x$. Hence

$$\begin{aligned} p(\text{observed death}) &= \int_0^1 [1 - H(x)]dF(x) \\ &= F(1) - \int_0^1 H(x)dF(x) \end{aligned} \tag{1}$$

demonstrating that the true death rate exceeds the chance of observed death by a nonnegative integral. Similarly

$$p(\text{observed withdrawal}) = \int_0^1 [1 - F(x)]dH(x) \tag{2}$$

$$= H(1) - \int_0^1 F(x)dH(x) \ .$$

To survive he must neither die nor withdraw, hence

$$p(\text{survival}) = [1 - F(1)][1 - H(1)] \ . \tag{3}$$

It is easy to verify that these probabilities add to one. (This provides an unusual proof of the usual formula for integration by parts!) To economize on notation we write $F(1) = \phi$ , $H(1) = \theta$ and define

$$\rho = \int_0^1 F(x)dH(x)/F(1)H(1) \ . \tag{4}$$

A probabilistic interpretation of $\rho$ is that it is the conditional probability of death, given that both X and Y are less than one: that is, given that both withdrawal and death take place in the interval. It is an artifice of the model, rather than a physically interpretable quantity. Similarly $(1 - \rho)$ is the probability of withdrawal under the same conditions. Then

$$\begin{cases} p(\text{observed death}) = \phi\{1 - (1 - \rho)\theta\} \ , \\ p(\text{observed withdrawal}) = \theta(1 - \rho\phi) \ , \text{ and} \\ p(\text{survival}) = (1 - \phi)(1 - \theta) \ . \end{cases} \tag{5}$$

The likelihood for the data is consequently

$$\phi^D(1 - \rho\phi)^W(1 - \phi)^S\{1 - (1 - \rho)\theta\}^D\theta^W(1 - \theta)^S \ . \tag{6}$$

Within the restrictions of the model this expression contains all the information in the data and several important remarks can be made about it. First, although the conversation was extended from $F(1)$ and $H(1)$ to include these functions throughout the whole of the interval, in fact the only additional quantity thereby introduced is $\rho$ . Consequently the data only contains information on three aspects of $F$ and $H$ , namely $\phi$ , $\theta$ and $\rho$ . The value of $\rho$ will be considered below. Second, the trinomial likelihood (6) depends on these three parameters $\phi$ , $\theta$ and $\rho$ , each of which, irrespective of the values of the others, can assume any value in the unit interval. Consequently in the unit cube of $(\phi,\theta,\rho)$-values, there is a one-dimensional curve on which, as $\phi$ , $\theta$ and $\rho$ vary, the probabilities (5) remain unaltered, and we have a problem that econometricians describe as unidentified: that is, even if the probabilities on the left-hand sides of (5) were known exactly $(N \to \infty)$ the values of $\theta$ , $\phi$ and $\rho$ would not be known but only known to lie on the curve just mentioned. Third, if $\rho$ were known, then the likelihood (6), now dependent only on $\theta$ and $\phi$ , factors into a product of two terms which depend on $\theta$ alone and $\phi$ alone. Consequently estimation of each can proceed separately if $\theta$ and $\phi$ are judged independent. In particular, as some writers suppose, if the withdrawal times are deterministic then only the $\phi$-part is relevant. But with $\rho$ unknown, the situation is more complicated and there is no obvious reason for discarding the $\theta$-component, when intersected in $\phi$ , since it appears to contain information about $\rho$ which may be of value in interpreting the $\phi$-component. Hence $\rho$ , which arises from the grouping, plays a key role in the analysis, and we now discuss it.

Since $F$ and $H$ are both increasing, $0 \leq \rho \leq 1$. Both extremes are attainable: for if all withdrawals take place at the beginning (end) of the interval, but no deaths occur there, then $\rho = 0(1)$. If linear splines are used to approximate $F$ and $H$, so that they are both linear in $(0,1]$, then simple calculation shows $\rho = \frac{1}{2}$. More generally, if $F(x) = cH(x)$, then $\rho = \frac{1}{2}$, for any $c$: in particular if $c = 1$ and they are identical. If deaths and withdrawals take place at constant rates, $\lambda$ and $\mu$ respectively, during the interval, so that $F(x) = 1 - e^{-\lambda x}$ and $H(x) = 1 - e^{-\mu x}$, then simple calculation shows that

$$\rho = \frac{\lambda - (\lambda + \mu)e^{-\mu} + \mu e^{-(\lambda+\mu)}}{(\lambda + \mu)(1 - e^{-\lambda})(1 - e^{-\mu})} . \tag{7}$$

Expanding both numerator and denominator in powers of $\lambda$ and $\mu$ up to and including *third* powers shows that $\rho \approx \frac{1}{2}$ for all $\lambda$ and $\mu$. Hence unless the death- and withdrawal-rates are high, $\rho = \frac{1}{2}$ is a good approximation. Other assumptions are possible: for example, Chiang (1961) supposes that all withdrawals take place at the mid-point of the interval and that the deaths occur at a constant rate. Then $\rho = F(\frac{1}{2})/F(1) = \{1 - (1 - \phi)^{\frac{1}{2}}\}\phi^{-1}$.

These considerations, allied to the unidentifiability aspect, mean that two types of procedure are possible. First, assume enough about $F$ and $H$ for $\rho$ either to be known or to be known as a function of $\phi$ and $\theta$. (Thus linearity gives $\rho = \frac{1}{2}$ : Chiang's procedure is an example of the second possibility.) Second, leave $F$ and $H$ sufficiently general for $\rho$ to be unknown, even as functions of $\phi$ and $\theta$, and to face up to the identifiability problem. Previous work on the problem has concentrated on the first procedure, making assumptions about $F$ and $G$. However Breslow and Crowley (1974) show that if this is done then the resulting estimators of $\phi$ and $\theta$ are typically inconsistent, in the sense that as the size of the data base increases, $N \to \infty$, they do not tend to the true death- and withdrawal-rates. For example, the classical estimate, $D/(N - \frac{1}{2}W)$ is only consistent under rather special assumptions about $F$ and $H$. In this paper we therefore take the second view, that $\rho$ is not known, even as a function of $\phi$ and $\theta$. We show that estimation is still possible and that the possible inconsistency is avoided.

We therefore return to the likelihood, given by (6), in which $\phi$, $\theta$ and $\rho$ can take all values in the unit cube. Since $\rho$ is unknown it has a distribution, which we suppose independent of $\phi$ and $\theta$. Let its density be $f(\rho)$. (These assumptions will be discussed below when we are in a better position to appreciate the roles they play.) Consequently, from (6)

$$(8) \quad p(D,W,S \mid \theta,\phi) = \phi^D(1-\phi)^S\theta^W(1-\theta)^S \int_0^1 (1-\rho\theta)^W\{1-(1-\rho)\theta\}^D f(\rho)d\rho \; .$$

Even with simple forms for $f(\rho)$, for example Beta, the integral cannot be evaluated in terms of standard functions. There are two possible procedures: for numerical values of $D$, $S$ and $W$ to use numerical integration techniques; for large $N$ to use analytic approximations. In the present paper we explore the latter possibility. This will give us the "feel" of the situation more easily than the computational method, though that method will be essential for small samples.

If $N$, and consequently $D$ and $W$, are large, the integrand in (8), as a function of $\rho$, has a maximum where

$$\frac{1 - (1 - \rho)\theta}{D\theta} = \frac{1 - \rho\phi}{W\phi}$$

or

$$\rho = \frac{d}{\phi} - \frac{w(1 - \theta)}{\theta} = \eta \text{ , say} \tag{9}$$

where $d = D/(D + W)$ and $w = W/(D + W)$ ; $d + w = 1$ . (The interpretation of $d$ is the proportion of those seen to die out of all who left the system in the interval.) This value of $\rho$ , that has been written $\eta = \eta(\phi,\theta)$ , is the maximum likelihood value of $\rho$ for given $\phi$ and $\theta$ .

There are two possibilities: if $\eta$ lies outside the range of integration in (8), namely the unit interval, then the integrand will be small; if $0 < \eta < 1$ then the integrand will be appreciable only in the neighborhood of the maximum. We shall therefore suppose the likelihood vanishes outside $0 < \eta < 1$ .

We next use the result that if $0 < \eta < 1$

$$\int_0^1 e^{NL(\rho)} f(\rho) d\rho \sim e^{NL(\eta)} f(\eta)\{-L''(\eta)\}^{-\frac{1}{2}} \tag{10}$$

asymptotically as $N \to \infty$ . (The double prime denotes a second derivative.) This is easily established by a steepest-descent type of argument. On applying this to (8) with

$$NL(\rho) = W \log (1 - \rho\phi) + D \log \{1 - (1 - \rho)\theta\}$$

and simplifying the result, we have for $p(D,W,S \mid \theta,\phi)$

$$(1 - \phi)^S (1 - \theta)^S [1 - (1 - \theta)(1 - \phi)]^{N-S+1} f(\eta)\theta^{-1}\phi^{-1} .$$

A change of variables from $\phi$ and $\theta$ to $\zeta = (1 - \theta)(1 - \phi)$ and $\eta$ gives

$$(11) \qquad p(D,W,S \mid \zeta,\eta) = \zeta^S(1 - \zeta)^{N-S} f(\eta) \cdot (1 - \zeta)/\theta\phi \ .$$

It may help in the appreciation of what is happening to consider $p(D,W,S \mid \zeta,\eta)$ in the plane of $\phi$ and $\theta$ (see the figure). It is zero (or strictly, very small) except between the two curves $\eta = 0$ and $\eta = 1$ drawn in the figure. Between these two curves the function is only large in the neighborhood of the curve $\zeta = S/N$, also drawn in the figure. Consequently we need only concern ourselves with this probability in a region indicated by hatching in the figure. Within this region and along curves $\zeta$ = constant, the probability is dominated by $f(\eta)$ where $f$ is the prior density for $\rho$. Furthermore if the final term in (11), $(1 - \zeta)/\theta\phi$, which comes from the second derivative in (10), is ignored, the probability (or more correctly, likelihood) factors into functions of $\zeta$ and $\eta$ separately. The curve $\zeta = S/N$ intersects the two curves $\eta = 0$ and $\eta = 1$ at A and B in the figure, whose $\phi$-coordinates are respectively $D/N$ and $D/(N - W)$. Hence with probability almost one, $\phi$ lies between these two values which corresponds to assuming that all withdrawals take place at the end (beginning) of the interval. A common estimate is to adopt the compromise $D/(N - \frac{1}{2}W)$. Most of these results can be deduced by equating the chances in (5) to $D/N$, $W/N$ and $S/N$ respectively. The additional feature that our analysis shows is the factorization of the likelihood when expressed in terms of $\zeta$ and $\eta$, and the explicit form of the factors, in particular that in $\eta$, which is entirely due to the prior opinions about $\rho$, $f(\cdot)$.

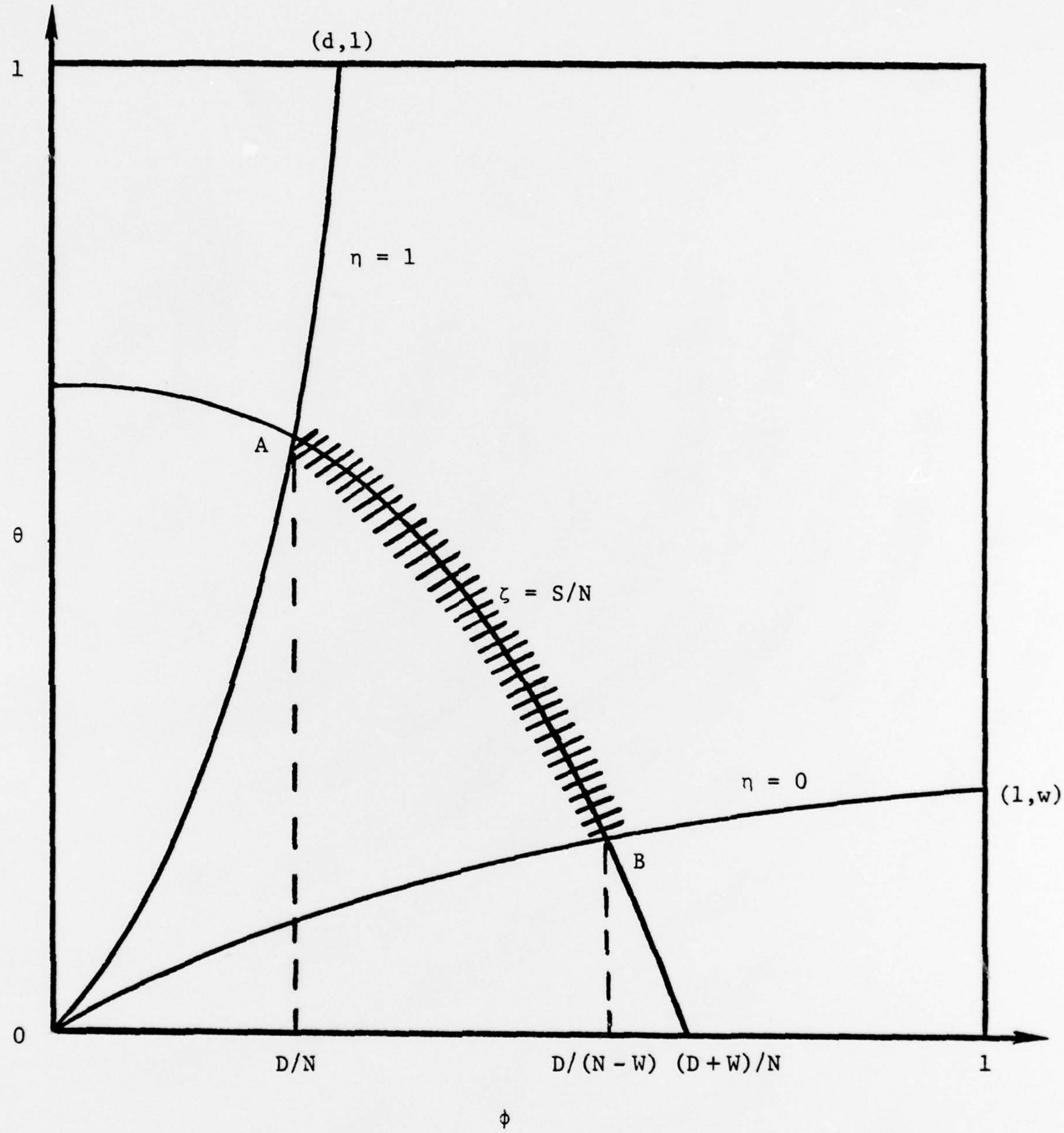



FIGURE 1

If, over the hatched region, the prior density for $\phi$ and $\theta$ is reasonably constant then the expression on the right-hand side of (11) is approximately proportional to the probability of $\zeta$ and $\eta$, given $D$, $W$ and $S$. If the contribution from the second derivative is ignored, we have approximately

$$p(\zeta,\eta \mid D,W,S) \propto \zeta^{S}(1-\zeta)^{N-S} f(\eta) \tag{12}$$

so that $\zeta$ and $\eta$ are independent, $\zeta$ having a Beta-distribution and $\eta$ having the density of $\rho$. Admittedly rather crude approximations have been made in order to obtain this result, but experience with steepest-descent type arguments in this sort of situation suggests that they will work reasonably well and not demand excessively large data sets (large values of $N$). They are about as good as maximum likelihood results are.

The posterior mode of (12) occurs at $\zeta = s$ and $\eta = r$, where $s = S/N$, the proportion who survive the interval, and $r$ is the most probable a priori value of $\rho$. Translating this back into values of $\phi$ and $\theta$, the modes satisfy

$$\begin{cases} (1-\phi)(1-\theta) = s \quad \text{and} \\ \dfrac{d}{\phi} - \dfrac{w(1-\theta)}{\theta} = r\,. \end{cases} \tag{13}$$

Solving for $\phi$, the chance of death in the interval, we easily obtain the quadratic equation

$$\phi^2 r - \phi\{(d+r) + s(w-r)\} + d(1-s) = 0\,. \tag{14}$$

In terms of the raw data $D$, $W$ and $N$, this is

$$\phi^2 rN - \phi\{N - W + r(D + W)\} + D = 0 \ . \tag{15}$$

It was remarked in the third comment after (6) that if $\rho$ is known the $\phi$-part of the likelihood is

$$\phi^D (1 - \rho\phi)^W (1 - \phi)^S \ .$$

It is easy to verify that the maximum of this likelihood satisfies (15) with $\rho = r$. In other words, our posterior mode for $\phi$ is the maximum of the $\phi$-part of the likelihood with $\rho$ replaced by its most likely value, $r$. The left-hand side of (14) equals $d(1 - s) > 0$ at $\phi = 0$ and equals $-s(1 - r) < 0$ at $\phi = 1$, so that there exists just one root in the unit interval. At the value of $\phi$ given by

$$\phi_1 = \frac{D}{N - (1 - r)W} \tag{16}$$

the left-hand side of (15) is easily calculated to be $r(1 - r)D^2W/N^3$, which is small. Hence in many cases $\phi_1$ is a reasonable estimate of the death-rate. It can be justified on intuitive grounds by supposing that all withdrawals take place at $r$, so that each withdrawal is only exposed to the risk of "observed" death for time $r$. If this happens then $\rho = r$. A usual assumption is that $r = \frac{1}{2}$, which we have seen is a reasonable value of $\rho$. It is perhaps worth noting that at $\phi_1$ the quadratic is positive so that $\phi_1$ is always less than the posterior mode, hence there may be a systematic error in using $\phi_1$ in under-estimating the death-rate.

Further progress may be made by returning to the density of $\zeta$ and $\eta$, Equation (12), and considering the second moments. Clearly $\zeta$ has approximate variance $s(1 - s)/N$, $\eta$ has variance $v$, say, the prior variance of $\rho$, and they are uncorrelated. From these second moments of $\zeta$ and $\eta$ we can pass to the approximate second moment of $\phi$ using the delta-method. The expression for $\phi$ in terms of $\zeta$ and $\eta$ is (Equation (14) with $\zeta$ for $s$, and $\eta$ for $r$)

$$\phi^2\eta - \phi\{d + \eta + \zeta(w - \eta)\} + d(1 - \zeta) = 0 .$$

Hence

$$\begin{aligned} &2\phi\eta\Delta\phi + \phi^2\Delta\eta - \{d + \eta + \zeta(w - \eta)\}\Delta\phi \\ &- \phi\{\Delta\eta + (w - \eta)\Delta\zeta - \zeta\Delta\eta\} - d\Delta\zeta = 0 , \end{aligned}$$

and

$$\Delta\phi = \frac{\{d + \phi(w - \eta)\}\Delta\zeta + \{\phi(1 - \zeta) - \phi^2\}\Delta\eta}{2\phi\eta - \{d + \eta + \zeta(w - \eta)\}} .$$

Consequently

$$(17) \quad \mathrm{var}\,(\phi \mid d,s,N) \approx \frac{\{d + \hat{\phi}(w - r)\}^2 s(1 - s)/N + \hat{\phi}^2\{1 - s - \hat{\phi}\}^2 v}{\{2\hat{\phi}r - d - r - s(w - r)\}^2}$$

where $\hat{\phi}$ is the solution of (14). (Note: $d + w = 1$ .) The form of this variance is particularly interesting. It consists of two parts, derived respectively from the variances of $\zeta$ and $\eta$, the first of which tends to zero as $N \to \infty$, but the second of which does not (unless the proportion of deaths or withdrawals observed tends to zero).

The latter part is a multiple of $v$, the prior variance of $\rho$. Consequently, however large $N$ is, there remains some uncertainty as to the true death-rate, ascribable to the original uncertainty about $\rho$. The estimator given by (14) is therefore not inconsistent as suggested by Breslow and Crowley. Neither is it consistent. There is just not enough information in the data, however much of it there is, to be sure about the death-rate. The original uncertainty about $\rho$ is present however large is the data set.

It is of some interest to see the effect of the prior variance of $\rho$, $v$, on the very large sample variance of $\phi$. Letting $N \to \infty$ in (17), approximating $\hat{\phi}$ by $\phi$, $r$ by $\rho$ and using the facts that $D/N$ and $S/N$ tend to the chances given in (5), tedious algebra establishes that the variance of $\phi$ tends to

$$\left\{\frac{\phi(1-\phi)\theta}{(1-\theta)+\rho(\theta-\phi)}\right\}^2 v \, .$$

For example, if $\phi = \theta$ so that the withdrawal- and death-rates are equal, this is $\phi^2 v$, or the standard deviation is scaled down from $v^{\frac{1}{2}}$ for $\rho$ to $\phi v^{\frac{1}{2}}$ for $\phi$. If, as we suggest below, $v$ is small, the residual uncertainty about $\phi$ is even smaller, especially for low death-rates.

Discussion:

Many statisticians will object to our argument because it introduces a prior distribution for $\rho$ , and, less importantly, $\phi$ and $\theta$ . We will discuss this in a moment, but first I would ask them to consider the operational consequences of the method. Rather than committing oneself to a rigid assumption about F and H , the method allows for flexibility in these functions and does not involve a committment to any particular form. The result of this is that instead of producing estimators which are inconsistent unless the rigid assumption is exactly true, information is provided about the uncertainty in the death-rate and the role played in this uncertainty by the lack of knowledge of $\rho$ . Our method therefore seems both flexible and realistic.

Here is no place to argue in favor of $\rho$ having a distribution. Essentially without it inconsistency of judgments cannot be avoided. The choice of the exact form of f is not a subject that a statistician should consider on his own: he must necessarily consult with the demographer or reliability engineer who is familiar with the situation. The statistician's task is to help the scientist to articulate his knowledge of the problem in probabilistic terms. Here it may not be unreasonable to assume $\rho$ has a fairly tight distribution around $\rho = \frac{1}{2}$ , since, as we saw above, many familiar situations give values of $\rho$ either exactly $\frac{1}{2}$ or very near to it; and $\rho$ near 0 or 1 only arises in extreme situations. But there is one case where $\rho$ is near 3/4 , namely in calendar-year counts of academic staff, since most withdrawals take place at the end of the academic year in September. In our large-sample analytic treatment only the mode and variance of $\rho$ matter. With a uniform distribution of $\rho$ , $v = 1/12$ , and values substantially less than this seem in order.

We have made the severe assumption that $\rho$ is independent of $\phi$ and $\theta$. The appropriateness of this must be a matter for discussion and cannot be decided on purely statistical grounds. It would not be difficult to incorporate dependence into the argument but without a specific application in mind it seems idle to speculate on the form it might take. In some cases $\rho$ might be a function of $\theta$ and $\phi$, as when $F$ and $H$ are both exponential (Equation (7) above), identifiability is ensured and consistency obtainable though the likelihood is rather complicated.

We have considered only a single interval. For several intervals $I_k$, it is easy to see the likelihood is the product of factors like (6) with data $D_k$, $W_k$, $S_k$ $(S_{k-1} = D_k + W_k + S_k)$ and parameters $\phi_k$, $\theta_k$, $\rho_k$. Consequently the likelihood factors. However it seems reasonable that the priors will *not* factor since the death-rates $\phi_k$, for example, are likely to be similar in adjacent intervals. This correlation will lead to smoothing of the rates as a function of $k$, but details remain to be worked out.

I am grateful to Professor Richard E. Barlow who invited me to Berkeley, introduced me to life-table problems and challenged me to provide a Bayesian analysis. This paper is the reply to the challenge.

## REFERENCES

Breslow, N. and J. Crowley (1974), "A Large Sample Study of the Life Table and Product Limit Estimates Under Random Censorship," Annals of Statistics, Vol. 2, pp. 437-453.

Chiang, C. L. (1961), "A Stochastic Study of the Life Table and its Applications III," Biometrics, Vol. 17, pp. 57-78.